

ANALÍZIS II.

Bártfai Pál

11. Kétfváltozós függvények

11.1. Deriválás

A $z = f(x, y)$ kétfváltozós függvényénél a z függő változó értékét az x és az y független változók értékéből számoljuk ki. A függvényt háromdimenziós koordinátarendszerben tudjuk ábrázolni, ahol a függvény értelmezési tartománya az x, y -koordinátasíkon helyezkedik el, és az összetartozó (x, y, z) értékekhez rendelt pontok az értelmezési tartomány felett elhelyezkedő felületet képezik.

D. Az (a, b) pont δ sugarú S környezetét azok az (x, y) pontok alkotják, amelyeknek az (a, b) -től mért távolsága δ -nál kisebb, képletben $S = \{(x, y): (x - a)^2 + (y - b)^2 < \delta^2\}$. Az $f(x, y)$ függvényt az (a, b) pontban *folytonosnak* nevezzük, ha $\forall \varepsilon > 0$ -hoz az (a, b) pontnak van olyan S környezete, hogy $|f(x, y) - f(a, b)| < \varepsilon$, ha $(x, y) \in S$.

A kétfváltozós differenciálhatóság definíciójához előbb fogalmazzuk át az egyváltozós definíciót. $f(x)$ differenciálható az a pontban, ha a megváltozása $f(a + h) - f(a) = Ah + r(h)h$ alakban írható fel, ahol $r(h) \rightarrow 0$, ha $h \rightarrow 0$. Az A értéket a függvény a pontbeli deriváltjának neveztük. Korábban ezt a definíciót h -val elosztott alakban használtuk.

D. Az $f(x, y)$ függvény az (a, b) pontban *differenciálható*, ha megváltozása

$$f(a + h, b + k) - f(a, b) = Ah + Bk + r_1(h, k)h + r_2(h, k)k$$

alakban írható fel, ahol $r_1(h, k) \rightarrow 0$ és $r_2(h, k) \rightarrow 0$, ha a két pont távolsága $\sqrt{h^2 + k^2} \rightarrow 0$.

Az A és a B számok az $f(x, y)$ *parciális deriváltjai* az (a, b) pontban. A parciális deriváltak jelölése: $f'_x(a, b) = A$, ill. $f'_y(a, b) = B$. (Használatos még a $\frac{df}{dx}$ -nek megfelelő $\frac{\partial f}{\partial x}$, ill. $\frac{\partial f}{\partial y}$

jelölés is.) A fenti definícióban $k = 0$ helyettesítéssel

$$f(a + h, b) - f(a, b) = Ah + r_1(h, k)h,$$

azaz

$$\frac{f(a + h, b) - f(a, b)}{h} \rightarrow A = f'_x(a, b)$$

adódik, vagyis az x szerinti parciális derivált az y érték rögzítésével kapott egy változós függvény deriváltja. Hasonlóan az x érték rögzítése mellett végezve el a differenciálást, az y szerinti parciális deriváltat kapjuk.

Ha az y értékét rögzítjük, legyen $y = b$, akkor a felület $y = b$ által meghatározott metszetéről van szó. Az x szerinti parciális derivált tehát a felület x -tengellyel párhuzamos metszetgörbéjének a deriváltja (metszetgörbét itt és a következőkben - ha mást nem mondunk - mindig a z tengellyel párhuzamos síkkal képezzük). A parciális deriváltak létezhetnek (a metszetgörbék differenciálhatók lehetnek) anélkül, hogy a kétfváltozós függvény differenciálható lenne.

11.2. Összetett függvény deriváltja

T. Ha az $u(t)$ és a $v(t)$ függvények differenciálhatók a $t = a$ pontban, és $f(x, y)$ is differenciálható az $(u(a), v(a))$ pontban, akkor

$$(f(u(t), v(t)))' = f'_x(u(t), v(t))u'(t) + f'_y(u(t), v(t))v'(t).$$

B. Írjuk fel az összetett függvény differenciálhatóságára vonatkozó képletet, és használjuk benne a $\Delta u = u(a+h) - u(a)$, $\Delta v = v(a+h) - v(a)$ rövidebb jelöléseket:

$$f(u(a+h), v(a+h)) - f(u(a), v(a)) = f'_x(u(a), v(a))\Delta u + f'_y(u(a), v(a))\Delta v + r_1\Delta u + r_2\Delta v,$$

ahol $r_i = r_i(\Delta u, \Delta v) \rightarrow 0$, ha $(\Delta u)^2 + (\Delta v)^2 \rightarrow 0$ ($i = 1, 2$). Osszuk el h -val mindkét oldalt, és végezzük el a $h \rightarrow 0$ határátmenetet. A baloldal $f(u(t), v(t))$ $t = a$ pontbeli deriváltja lesz, míg a jobb oldalon $\frac{\Delta u}{h} \rightarrow u'(a)$ és $\frac{\Delta v}{h} \rightarrow v'(a)$, ami a bizonyítandó állítást adja.

11.3. Iránymenti derivált

A parciális deriváltak az x - ill. az y -tengelyekkel párhuzamos metszetgörbék iránytangensét adják meg. Érdekes lehet azonban más irányú metszetek iránytangense is. Például, ha a felület egy táj domborzatát szemlélteti, akkor kérdezhetjük az ÉK-i irányban haladó ösvény meredekségét, a vízszintesen haladó ösvény irányát, a labda legurulásának az irányát. Ezek a feladatok az iránymenti derivált kiszámításával oldhatók meg.

Tegyük fel a fejezetben, hogy a kétváltozós függvényünk differenciálható. Az (a, b) pontban adjunk meg egy irányt, mely α szöveget zár be az x -tengellyel ($-\frac{\pi}{2} < \alpha \leq \frac{\pi}{2}$). Írjuk fel az x -tengellyel α szöveget bezáró síkkal képezett metszetgörbe egyenletét vigyázva arra, hogy léptékhelyes maradjon. Az (a, b) ponttól t távolságra és α irányban lévő pont koordinátái $(a + t \cos \alpha, b + t \sin \alpha)$, vagyis a metszetgörbe egyenlete $f(a + t \cos \alpha, b + t \sin \alpha)$. Ennek t szerinti deriváltja az összetett függvény deriválási szabálya szerint:

$$\frac{d}{dt} f(a + t \cos \alpha, b + t \sin \alpha) = f'_x(a + t \cos \alpha, b + t \sin \alpha) \cos \alpha + f'_y(a + t \cos \alpha, b + t \sin \alpha) \sin \alpha$$

ami $t = 0$ helyettesítéssel adja a metszetgörbe deriváltját az (a, b) pontban.

D. A kétváltozós $f(x, y)$ függvény (a, b) pontbeli, α irányú iránymenti deriváltjának nevezzük a függvény adott irányú megváltozásával képezett különbségi hányados határértékét.

T. Az $f(x, y)$ függvény (a, b) pontbeli, α irányú iránymenti deriváltja az

$$f'_x(a, b) \cos \alpha + f'_y(a, b) \sin \alpha$$

képlettel számolható ki a parciális deriváltak segítségével.

11.4. Magasabbrendű parciális deriváltak

D. Az $f(x, y)$ függvényt kétszer differenciálhatónak nevezzük az (a, b) pontban, ha a parciális deriváltjai a pont környezetében léteznek és az (a, b) pontban differenciálható függvények.

Az $f(x, y)$ függvénynek összesen négy másodrendű parciális deriváltja van: differenciálhatok kétszer x szerint, vagy először x szerint, majd y szerint, vagy ugyanezt tehetem, csak fordított sorrendben, végül kétszer y szerint. Az egyes parciális deriváltak jelölése: $f''_{xx}, f''_{xy}, f''_{yx}, f''_{yy}$.

T (Young tétele). Ha $f(x, y)$ az (a, b) pont valamely S környezetében kétszer differenciálható, és ott a parciális deriváltak folytonosak, akkor ott $f''_{xy} = f''_{yx}$.

M. Young eredeti tétele az $f(x, y)$ az (a, b) pontban való kétszeri differenciálhatósága esetén bizonyítja ugyanezt az állítást.

B. Képezzük a $D(h) = f(a + h, b + h) - f(a + h, b) - f(a, b + h) + f(a, b)$ kifejezést. (Ezt az f függvény megváltozásának is nevezik a szóban forgó csúcspontokkal megadott téglalapon.) Ha bevezetjük az $u(x) = f(x, b + h) - f(x, b)$ függvényt, és alkalmazzuk kétszer a Lagrange-féle középérték tételt, akkor

$$D(h) = u(a + h) - u(a) = u'(\xi)h = (f'_x(\xi, b + h) - f'_x(\xi, b))h = f''_{xy}(\xi, \eta)h^2,$$

amiből

$$\lim_{h \rightarrow 0} \frac{D(h)}{h^2} = f''_{xy}(a, b).$$

Hasonlóan $v(x) = f(a + h, x) - f(a, x)$ jelöléssel

$$D(h) = v(b + h) - v(b) = v'(\eta)h = (f'_y(a + h, \eta) - f'_y(a, \eta))h = f''_{yx}(\xi, \eta)h^2,$$

amiből

$$\lim_{h \rightarrow 0} \frac{D(h)}{h^2} = f''_{yx}(a, b),$$

ami a tétel állítását jelenti.

11.5. Kétváltozós lokális szélsőérték

D. Az $f(x, y)$ függvénynek az (a, b) pontban *lokális maximuma (minimuma)* van, ha (a, b) -nek van olyan környezete, amelyben $f(a, b)$ a maximális (minimális) függvényérték. A lokális maximum és a lokális minimum közös neve: *lokális szélsőérték*.

ST. Ha $a > 0$ és $b^2 < ac$, akkor az $g(x) = a \cos^2 x + 2b \sin x \cos x + c \sin^2 x$ függvény minden helyettesítési értéke pozitív.

B. Teljes négyzetté történő kiegészítéssel

$$f(x) = a \left(\cos x + \frac{b}{a} \sin x \right)^2 + \frac{ac - b^2}{a} \sin^2 x \geq 0,$$

és nulla csak akkor lehetne, ha mindkét nemnegatív tag nulla, a második azonban csak $\sin x = 0$ esetén válik nullává, de ekkor az első tag nem nulla.

M. Itt és a következő tételben az $ac - b^2$ kifejezés a megadott másodfokú kifejezés diszkriminánsa (pontosabban 4-gyel elosztva), ami az $a > 0$ feltétellel párosítva a másodfokú kifejezés pozitivitását biztosítja.

T. Az $f(x, y)$ függvényről tételezzük fel, hogy az (a, b) pont valamely S környezetében kétszer differenciálható, és ott a másodrendű parciális deriváltak folytonosak. Ha $f'_x(a, b) = f'_y(a, b) = 0$, és

$$f''_{xx}(a, b)f''_{yy}(a, b) > (f''_{xy}(a, b))^2,$$

akkor a függvénynek lokális szélsőértéke van az (a, b) pontban. Ha $f''_{xx}(a, b) > 0$, akkor a szélsőérték lokális minimum, ha $f''_{xx}(a, b) < 0$, akkor lokális maximum.

B. Feltehetjük, hogy $(a, b) = (0, 0)$, ugyanis a függvény eltolható úgy, hogy az (a, b) pontnak az origó feleljen meg. Ez a transzformáció a deriváltak értékét és a lokális szélsőérték meglétét nem változtatja meg.

Vezessük be a $g(t) = f(t\cos\alpha, t\sin\alpha)$ jelölést, és fejtsük Taylor-sorba két tagig Lagrange-féle maradéktaggal ellátva:

$$g(t) = g(0) + g'(0)t + \frac{g''(t_0)}{2}t^2.$$

ahol $|t_0| < |t|$. $g'(0)$ az iránymenti derivált, és mivel az első parciális deriváltak nullák, $g'(0) = 0$, így

$$g(t) - g(0) = \frac{t^2}{2} [(g''(t_0) - g''(0)) + g''(0)].$$

$\xi = t_0\cos\alpha$, $\eta = t_0\sin\alpha$ jelöléssel élve

$$g''(t) = f''_{xx}(\xi, \eta)\cos^2\alpha + 2f''_{xy}(\xi, \eta)\cos\alpha\sin\alpha + f''_{yy}(\xi, \eta)\sin^2\alpha,$$

ebből a $|\sin\alpha| \leq 1$ és $|\cos\alpha| \leq 1$ egyenlőtlenségeket és a második deriváltak folytonosságát felhasználva

$$|g''(t_0) - g''(0)| \leq |f''_{xx}(\xi, \eta) - f''_{xx}(0, 0)| + 2|f''_{xy}(\xi, \eta) - f''_{xy}(0, 0)| + |f''_{yy}(\xi, \eta) - f''_{yy}(0, 0)| < 4\varepsilon,$$

ha $|t_0| < \delta$. Másrészt $g''(0) = f''_{xx}(0, 0)\cos^2\alpha + 2f''_{xy}(0, 0)\cos\alpha\sin\alpha + f''_{yy}(0, 0)\sin^2\alpha$ α -nak folytonos függvénye, ezért felveszi a $[0, 2\pi]$ -ben az m -mel jelölt minimumát, de a segédétel miatt $m > 0$. Ha 4ε -t kisebbre választjuk, mint m , és a δ -t ehhez az ε -hoz választjuk, akkor $|t| < \delta$ esetén $g(t) - g(0) = f(x, y) - f(0, 0) > 0$, vagyis az origó δ -sugarú környezetében minimuma van $f(x, y)$ -nak.

M. Ha a differenciálhatósági és folytonossági feltételek teljesülnek, továbbá $f'_x(a, b) = f'_y(a, b) = 0$ és $f''_{xx}(a, b)f''_{yy}(a, b) < (f''_{xy}(a, b))^2$, akkor a függvénynek nincs szélsőértéke. Ha az utóbbi egyenlőtlenségben egyenlőség áll fenn, akkor a kritérium nem dönti el a kérdést.

11.6. Feltételes szélsőérték

Ha az $f(x, y)$ szélsőértékét azokra az (x, y) pontokra keressük, amelyekre a $\varphi(x, y) = 0$ feltétel teljesül, akkor az f feltételes szélsőértékéről beszélünk. A feltételes szélsőérték is - értelemszerűen - lehet globális és lokális szélsőérték. Ha az f felületet földrajzi domborzatnak képzeljük el, és a $\varphi(x, y) = 0$ a térképen (x, y) -síkra vett vetületben) egy ösvényt határoz meg, akkor a globális maximum az ösvény legmagasabb pontja, a lokális maximum a pont megfelelő környezetét tekintve az ösvény legmagasabb pontja.

A feltételes szélsőérték visszavezethető feltétel nélkülivé, ha $\varphi(x, y) = 0$ -ból valamelyik változót kifejezve az $f(x, y)$ -ba beírjuk és a kapott egyváltozós függvénynek számítjuk ki a feltétel nélküli - szélsőértékét. A változó kifejezése azonban nehézséget okozhat, ezért közvetlen megoldást keresünk.

T (Lagrange multiplikátor módszer). Képezzük az $L(x, y; \lambda) = f(x, y) - \lambda\varphi(x, y)$ kifejezést. Ha valamely $x_0, y_0; \lambda$ számhármásra $\varphi(x_0, y_0) = 0$, és a λ rögzített értéke mellett az $L(x, y; \lambda)$ kétváltozós függvénynek lokális szélsőértéke van az (x_0, y_0) pontban, akkor ebben a pontban az $f(x, y)$ -nak lokális feltételes szélsőértéke van a $\varphi(x, y) = 0$ feltétel mellett.

B. Azokban az (x, y) pontokban, melyekre a feltétel teljesül, $L(x, y; \lambda) = f(x, y)$. Így, ha L -nek szélsőértéke van egy pontban a pont valamely környezetét tekintve, akkor a feltételt kielégítő pontokban is igaz ugyanez, de itt $L = f$, tehát f -nek is szélsőértéke van a környezetben a feltételt kielégítő pontokat tekintve.

A megoldás menete a következő: képezzük a Lagrange-féle L kifejezést, megoldjuk az

$$L'_x = 0, L'_y = 0, \varphi(x, y) = 0$$

egyenletrendszert (három egyenlet, három ismeretlen) és ellenőrizzük L szélsőértékének létezését a kapott pontban vagy pontokban.

M. Ha az egyenletrendszert kielégítő pontban L -nek nincs szélsőértéke, akkor a feltételes feladatnak még lehet, csak ezzel a módszerrel nem dönthető el a kérdés. Itt az eldöntési kritérium eléggé gyenge, ellentétben a feltétel nélküli feladattal.

12. Határozott integrál

12.1. A határozott integrál definíciója

Legyen $f(x)$ korlátos függvény az (a, b) véges intervallumon. A határozott integrállal az $f(x)$ görbe és az x -tengely (a, b) intervalluma által közrefogott területet számoljuk ki (a tengely alá eső területrészt negatív előjellel véve). A terület kiszámításának a módja téglalapokkal való közelítés, majd határérték képzése.

Készítsük el a terület alsó és felső közelítését a következőképpen. Legyen

$$x_0 = a < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

az (a, b) intervallum tetszőleges felosztása. Jelölje M_i az f függvény szuprémumát, m_i az f infimumát az (x_{i-1}, x_i) intervallumon.

D. Az integrál *felső közelítő összegének* az

$$S_n = \sum_{i=1}^n M_i (x_i - x_{i-1}),$$

alsó közelítő összegének az

$$s_n = \sum_{i=1}^n m_i (x_i - x_{i-1})$$

összeget nevezzük.

Ha a görbe alatti T terület létezik, akkor nyilván $s_n \leq T \leq S_n$.

A felosztás finomságát a leghosszabb részintervallum hosszával jellemezzük, vagyis a felosztás finomsága $\delta_n = \max(x_i - x_{i-1})$.

D. Ha S_n -nek és s_n -nek közös határértéke van, ha $\delta_n \rightarrow 0$, akkor az f az (a, b) intervallumon integrálható és a határérték az f határozott integrálja (a, b) -n. Jelölésben:

$$\lim_{\delta_n \rightarrow 0} s_n = \lim_{\delta_n \rightarrow 0} S_n = \int_a^b f(x) dx.$$

A definíció kissé pontosabban is elmondható: Ha $\forall \varepsilon > 0$ -hoz $\exists \delta$ és I , hogy $|S_n - I| < \varepsilon$ és $|s_n - I| < \varepsilon$, ha a felosztás δ -nál finomabb, akkor I az f határozott integrálja az (a, b) -n.

Ha nincs közös határérték, akkor azt mondjuk, hogy f nem integrálható.

(M. Az integrál fogalma többféleképpen is definiálható, a különböző integrálfogalmak kissé eltérhetnek egymástól. Ez a definíció Riemann-tól származik, ezért Riemann-integrálnak is nevezik.)

A következő két tételben az alábbi segédtelet felhasználjuk.

ST. Ha az f függvényre vonatkozóan $\forall \varepsilon > 0$ -ra $|S_n - s_n| < \varepsilon$, ha a felosztás δ -nál finomabb, akkor f integrálható.

B. Képezzük az összes lehetséges felső összeg infimumát, legyen ez I_1 , és az össze lehetséges alsó összeg szuprémumát, legyen ez I_2 . A tett feltétel szerint $|I_1 - I_2| < \varepsilon$, ahol ε tetszőleges, tehát $I_1 = I_2$. Mivel $s_n \leq I_1 \leq S_n$, a közelítő összegek I_1 -től való eltérése is ε -nál kisebb lesz, ami az integrálhatóságot jelenti.

T. Az $[a, b]$ -n folytonos függvény mindig integrálható.

B. Az $[a, b]$ -n folytonos f függvény egyenletesen is folytonos, azaz $\forall \varepsilon > 0 \exists \delta$, hogy $|f(x_2) - f(x_1)| < \varepsilon$, ha $x_1, x_2 \in [a, b]$ és $|x_2 - x_1| < \delta$. Legyen a felosztás finomsága δ , akkor

$$M_i - m_i = \sup_{\alpha, \beta \in [a, b]} f(\beta) - f(\alpha) \leq \varepsilon,$$

és ebből

$$0 \leq S_n - s_n = \sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) \leq \varepsilon \sum_{i=1}^n x_i - x_{i-1} = \varepsilon(b - a),$$

amiből látható, hogy a két közelítés eltérése tetszőlegesen kicsinnyé tehető, ha a felosztás finomságát elég kicsire választjuk.

T. Monoton függvény mindig integrálható.

B. Legyen a függvény pl. monoton növekedő. Az előzőhöz hasonlóan, ha a felosztás finomsága δ ,

$$0 \leq S_n - s_n = \sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) \leq \delta \sum_{i=1}^n M_i - m_i \leq \delta(f(b) - f(a)) \rightarrow 0.$$

A Riemann-integrál fogalma nem korlátos függvényre, ill. nem korlátos intervallumra nem értelmezhető, mert a közelítő összegek egyike végtelenné válik. Az integrál kiterjesztése ilyen esetre más úton történik, ezzel később foglalkozunk (lásd *improprius* integrál).

D. Az $\int_a^b f(x) dx$ integrál $a > b$ esetén is értelmezhető az $\int_a^b f(x) dx = -\int_b^a f(x) dx$ konvencióval.

T. Ha $a < c < b$, és f integrálható (a, b) -n, akkor $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$.

B. A felosztásban a c -t osztópontnak választva nyilvánvaló az állítás.

12.2. Az integrálfüggvény

D. Ha $f(x)$ integrálható (a, b) -n, akkor $x \in [a, b]$ esetén az

$$F(x) = \int_a^x f(u) du$$

függvényt *integrálfüggvénynek* nevezzük.

Az integrálfüggvény segítségével tetszőleges $(c, d) \subset (a, b)$ -re $\int_c^d f(x) dx = F(d) - F(c)$.

T. Ha $f(x)$ folytonos az x_0 pontban, akkor $F(x)$ differenciálható x_0 -ban, és $F'(x_0) = f(x_0)$.

B. Az előző pont utolsó tétele miatt

$$F(x_0 + h) - F(x_0) = \int_{x_0}^{x_0+h} f(u) du.$$

A különbségi hányados és a határérték eltérése egyetlen integrállal felírható, mert a konstans $f(x_0)$ integrálja téglalap módjára kiszámítható:

$$\frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) = \frac{1}{h} \int_{x_0}^{x_0+h} f(u) - f(x_0) du,$$

ebből

$$\left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right| \leq \frac{1}{h} \int_{x_0}^{x_0+h} |f(u) - f(x_0)| du \leq \varepsilon,$$

ha a h elég kicsi, ugyanis ekkor az f függvény folytonossága miatt az integrandus ε -nál kisebbé válik. A különbségi hányados határértéke tehát $f(x_0)$.

A határozott integrál kiszámításához tehát elég "kitalálni", hogy melyik az az F függvény amelynek a deriváltja f , ugyanis ekkor ez tekinthető integrálfüggvénynek.

12.3. Határozatlan integrál

D. Ha az $f(x)$ függvényhez található olyan $F(x)$ függvény, hogy F differenciálható (a, b) -n és $F'(x) = f(x) \forall x \in (a, b)$, akkor F -et az f *határozatlan integráljának*, vagy *primitív függvényének* nevezzük. A határozatlan integrál jelölésére az $\int f(x)dx$ jelölést használjuk.

Természetesen f határozatlan integrálja, ha egyáltalán létezik, nem egyértelmű, hiszen $(F(x) + c)' = F'(x)$, így $F(x)$ -szel együtt $F(x) + c$ is határozatlan integrálja f -nek.

T. f két határozatlan integrálja csak konstansban különbözhet. Ha megadjuk f határozatlan integrálját, akkor egy C állandót mindig hozzá kell adni feltüntetve, hogy a megadott függvény nem egyértelmű.

B. Legyen F_1 és F_2 a két határozatlan integrál, és legyen $G(x) = F_1(x) - F_2(x)$, akkor $G'(x) = f(x) - f(x) = 0 \forall x \in (a, b)$. Legyen $x_1 < x_2$, $x_1, x_2 \in (a, b)$ tetszőleges, akkor a Lagrange-féle középértéktétel miatt

$$G(x_1) - G(x_2) = G'(\xi)(x_1 - x_2) = 0,$$

vagyis $G(x_1) = G(x_2)$, tehát $G(x)$ konstans függvény.

T. Ha f folytonos $[a, b]$ -n, akkor van határozatlan integrálja.

B. Előző pont utolsó tételének állítása szerint nyilvánvaló.

M. Természetesen előfordulhat, hogy az integrálfüggvény létezik, de az nem differenciálható, így a határozatlan integrál már nem létezik.

12.4. Newton-Leibniz szabály

A primitív függvény (vagyis a határozatlan integrál) ismeretében a határozott integrál már kiszámítható.

T (Newton-Leibniz szabály). Ha f integrálható (a, b) -ben és primitív függvénye az $[a, b]$ intervallumon F , akkor $\int_a^b f(x)dx = F(b) - F(a)$.

B. Vegyük fel a (a, b) intervallumnak egy tetszőleges felosztását:

$$x_0 = a < x_1 < x_2 < \dots < x_{n-1} < x_n = b,$$

és jelöljük M_i -vel az f függvény szuprémumát, m_i -vel az f infimumát az (x_{i-1}, x_i) intervallumon, akkor a Lagrange-féle középértéktétel szerint

$$I = F(b) - F(a) = \sum_{i=1}^n F(x_i) - F(x_{i-1}) = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}).$$

Ebből kiindulva I -t két oldalról megbecsülhetjük:

$$s_n = \sum_{i=1}^n m_i(x_i - x_{i-1}) \leq I \leq \sum_{i=1}^n M_i(x_i - x_{i-1}) = S_n,$$

mivel f integrálhatóságát feltételeztük, S_n és s_n közös határértékhez, az $\int_a^b f(x)dx$ tart, ha a

felosztás finomsága tart nullához, ezért $I = \int_a^b f(x)dx$.

12.5. Alapintegrálok

A differenciálási táblázat mintájára elkészíthetjük az elemi függvények integráljait. A táblázat egyszerűen a deriválási táblázat megfordítása.

$f(x)$	$F(x)$	Érvényesség
x^α	$\frac{x^{\alpha+1}}{\alpha+1} + C$	$x > 0, \alpha \neq -1$
$\frac{1}{x}$	$\ln x + C$	$x > 0$
e^x	$e^x + C$	$x \in \mathbf{R}$
$\sin x$	$-\cos x + C$	$x \in \mathbf{R}$
$\cos x$	$\sin x + C$	$x \in \mathbf{R}$
$\frac{1}{1+x^2}$	$\arctg x + C$	$x \in \mathbf{R}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x + C$	$x \in [-1, 1]$
$\operatorname{sh} x$	$\operatorname{ch} x + C$	$x \in \mathbf{R}$
$\operatorname{ch} x$	$\operatorname{sh} x + C$	$x \in \mathbf{R}$

Míg az elemi függvényekből képezett képezett összegek, szorzatok, hányadosok és összetett függvények deriváltjai kiszámolhatók, képlettel megadhatók, addig az integrálok kiszámításánál a helyzet lényegesen rosszabb. Itt is vannak integrálási szabályok, de ezek sokkal korlátozottabban használhatók.

13. Az integrálás technikája

13.1. Elemi szabályok

T. Ha f_1 és f_2 integrálható függvények (a, b) -n, akkor $f_1 + f_2$ is integrálható és

$$\int_a^b f_1 + f_2 dx = \int_a^b f_1 dx + \int_a^b f_2 dx.$$

B. Vegyük fel a (a, b) intervallumnak egy tetszőleges felosztását

$$x_0 = a < x_1 < x_2 < \dots < x_{n-1} < x_n = b,$$

és jelöljük M_{1i} -vel és m_{1i} -vel az f_1 , M_{2i} -vel és m_{2i} -vel az f_2 szuprémumát ill. infimumát az i -edik intervallumon. Mivel $x \in (x_{i-1}, x_i)$ esetén $m_{1i} + m_{2i} \leq f_1(x) + f_2(x) \leq M_{1i} + M_{2i}$, az

$\int_a^b f_1 + f_2 dx$ felső összege kisebb, mint $\int_a^b f_1 dx$ felsőösszege meg $\int_a^b f_2 dx$ felső összege, de

ezek határértéke $\int_a^b f_1 dx + \int_a^b f_2 dx$. Ugyanezt kell elvégezni az alsó összegekkel is, akkor az

$\int_a^b f_1 + f_2 dx$ integrál alsó és felső összege ugyanazon határértékhez tart, tehát $f_1 + f_2$ integ-

rálható, és az integrál értéke $\int_a^b f_1 dx + \int_a^b f_2 dx$.

T. Ha f integrálható (a, b) -n, akkor cf is integrálható és $\int_a^b cfdx = c \int_a^b f dx$.

B. Ha $c > 0$, akkor cf felső (vagy alsó) közelítő összegéből a c kiemelhető és a másik tényező f felső (ill. alsó) közelítő összege lesz. Határátmenettel adódik az eredmény. Ha $c < 0$, akkor is kiemelhető c a cf felső (vagy alsó) közelítő összegéből, de a másik tényező f alsó (ill. felső) közelítő összege lesz. Határátmenettel ekkor is adódik az eredmény.

A határozatlan integrálokra vonatkozó hasonló állítások a deriválási szabályokból közvetlenül következnek.

13.2. Helyettesítéses integrálás

Az összetett függvény deriválási szabályának megfordított művelete.

T. Ha a g primitív függvénye G , és f differenciálható függvény, akkor $g(f(x)) \cdot f'(x)$ primitív függvénye $G(f(x))$. Más jelöléssel:

$$\int g(f(x))f'(x)dx = G(f(x)) + C.$$

B. Az összetett függvény deriválási szabálya alapján nyilvánvaló.

A tételt két formában alkalmazhatjuk integrálok kiszámítására. Az egyik lehetőség az, hogy észrevevessük, hogy az integrandus egy összetett függvény megszorozva a belső függvény deriváltjával, ekkor, ha a külső függvény határozatlan integrálját ismerjük, akkor az eredmény közvetlenül felírható.

P. Számítsuk ki az $\int \operatorname{tg} x dx$ integrált. Mivel $\operatorname{tg} x = -\frac{\sin x}{\cos x}$, és a $\cos x$ deriváltja $-\sin x$, az

$\frac{1}{\cos x}$ összetett függvény (a külső függvény az $1/x$, a belső a $\cos x$) szorozva van a belső függvény deriváltjával. Az $1/x$ primitív függvénye $\ln x$, így az eredmény:

$$\int \operatorname{tg} x dx = -\ln \cos x + C.$$

A másik lehetőség, amikor ezt közvetlenül nem ismerjük fel, de megkíséreljük az integrálban a $\cos x$ -et új változóval helyettesíteni, azaz legyen $y = \cos x$. Ekkor a $\frac{dy}{dx} = -\sin x$, vagyis a $\sin x dx$ helyére $-dy$ kerül:

$$\int \operatorname{tg} x dx = \int \frac{\sin x}{\cos x} dx = -\int \frac{1}{y} dy = -\ln y + C = -\ln \cos x + C.$$

A dx -szel történő "átszorzás" durva formalizmus, önmagában teljesen értelmetlen cselekedet, arra szolgál, hogy a fenti tételben szereplő eljárást mechanikussá tegyük.

A helyettesítési technika másik változatát egy újabb példán mutatjuk be, figyeljük meg a különbséget. Ezen a példán azt is látjuk, hogy az eljárás megismételhető, bár a kétszeri alkalmazás egy lépésben is elvégezhető, de nem mindig található meg az optimális helyettesítés egy lépésben.

P. Számítsuk ki az $\int \sqrt{1+\sqrt{x}} dx$ integrált. $y = \sqrt{x}$ helyettesítésnél $x = y^2$, $\frac{dx}{dy} = 2y$, azaz dx helyére $2y dy$ írandó, tehát

$$\int \sqrt{1+\sqrt{x}} dx = \int \sqrt{1+y} 2y dy.$$

Itt vesszük észre, hogy még az $1 + y$ -től is meg kell szabadulni, célszerű elvégezni a $z = \sqrt{1+y}$ helyettesítést. Mivel $y = z^2 - 1$, $dy = 2z dz$, így

$$\int \sqrt{1+\sqrt{x}} dx = \int \sqrt{1+y} 2y dy = \int z 2(z^2 - 1) 2z dz = 4 \int z^4 - z^2 dz = 4 \left(\frac{z^5}{5} - \frac{z^3}{3} \right) + C,$$

ahol z helyére vissza kell helyettesíteni a $z = \sqrt{1+\sqrt{x}}$ kifejezést.

A helyettesítés elvégzése után a határozatlan integrálnál mindig vissza kell térni az eredeti változóra! Ha határozott integrálról van szó, akkor akkor a visszatérés nem kötelező, de az integrálás határaival követni kell a helyettesítést. Ha az x szerinti integrálás az (a, b) intervallumra vonatkozik, és az $y = g(x)$ helyettesítést alkalmazzuk, akkor az y szerinti integrál határai $g(a)$ és $g(b)$.

P. Számítsuk ki az $\int_0^1 \sqrt{1+\sqrt{x}} dx$ integrált. Az eljárás ugyanaz, csak az első helyettesítés után

az $\int_0^1 \sqrt{1+y} 2y dy$ integrál adódik (a határok nem változnak, mert $\sqrt{0} = 0, \sqrt{1} = 1$). A második

helyettesítés után az

$$4 \int_1^{\sqrt{2}} z^4 - z^2 dz = 4 \left[\frac{z^5}{5} - \frac{z^3}{3} \right]_1^{\sqrt{2}} = 4 \left(\frac{\sqrt{2}^5}{5} - \frac{\sqrt{2}^3}{3} - \left(\frac{1}{5} - \frac{1}{3} \right) \right)$$

eredmény adódik. A primitív függvénynél szögletes zárójelben felül és alul tüntettük fel a határokat, ami azt jelenti, hogy a függvény felső határnál vett helyettesítési értékéből le kell vonni az alsó határnál kapott értéket (Newton-Leibniz szabály).

12.3. Parciális integrálás

A szorzat integrálási szabálya is átalakítható úgy, hogy integrálásnál felhasználható legyen.

T. Ha az $[a, b]$ intervallumban $u(x)$ és $v(x)$ differenciálható függvények és a deriváltjuk folytonos, akkor

$$\int u'v dx = uv - \int uv' dx$$

B. Mivel $(uv)'$ folytonos függvény $\int (uv)' dx = uv + C$. Írjuk fel a szorzat deriválási szabályát, $(uv)' = u'v + uv'$, és integráljuk mindkét oldalt

$$uv = \int (uv)' dx = \int u'v dx + \int uv' dx,$$

a C konstans nem tüntettük fel, mert a határozatlan integrál jelölés ezt a konstans magában foglalja.

P. Számítsuk ki az $\int xe^{-x} dx$ integrált. A "szereposztás" a következő: $u' = e^{-x}$, $v = x$. A parciális integrálást akkor alkalmazhatjuk, amikor az integrandus egyik tényezője könnyen integrálható, primitív függvénye nem bonyolult, a másik pedig a deriválással egyszerűsödik. A parciális integrálás elvégezhető:

$$\int xe^{-x} dx = -xe^{-x} + \int e^{-x} dx = -xe^{-x} - e^{-x} + C.$$

T. Határozott integrálra vonatkozó formula (az előző feltételek mellett):

$$\int_a^b u'v dx = [uv]_a^b - \int_a^b uv' dx.$$

12.4. Parciális törtekre bontás

Racionális törtfüggvények integrálásánál csak a legegyszerűbb esetet vesszük, amikor az n -edfokú nevezőnek n különböző valós gyöke van. A parciális törtek módszere más esetekben is hatásos, de a számolások nehezebbek.

Racionális törtfüggvény két polinom hányadosa. Ha a számláló fokszáma nem kisebb, mint a nevezőé, akkor egy polinom leválasztásával elérhető, hogy a törtkifejezésben a számláló már alacsonyabb fokú legyen. Ha a számláló $P_n(x)$, a nevező $Q_m(x)$ n - ill. m -edfokú polinomok ($n \geq m$), akkor $P_n(x) = A(x)Q_m(x) + B(x)$, ahol $B(x)$ fokszáma kisebb, mint m , így

$$\frac{P_n(x)}{Q_m(x)} = A(x) + \frac{B(x)}{Q_m(x)},$$

és az utóbbi törtben a számláló már alacsonyabb fokszámú.

Tegyük fel, hogy $Q_m(x)$ -nek m db. különböző valós gyöke van, legyenek ezek a_1, a_2, \dots, a_m , akkor feltehető, hogy $Q_m(x) = (x - a_1)(x - a_2)\dots(x - a_m)$.

T (parciális törtekre bontás). Ha $Q_m(x) = (x - a_1)(x - a_2)\dots(x - a_m)$, ahol az a_1, a_2, \dots, a_m számok különbözők és $P_n(x)$ m -nél alacsonyabb fokszámú polinom, akkor alkalmas A_1, A_2, \dots, A_m számokkal

$$\frac{P_n(x)}{Q_m(x)} = \sum_{k=1}^m \frac{A_k}{x-a_k}.$$

B. A bizonyítás teljes indukcióval történik. $m = 1$ -re az állítás triviális. Tegyük fel, hogy $m - 1$ fokszámú nevező esetén is igaz. Vonjunk le a baloldali törtből $\frac{A_m}{x-a_m}$ -et, akkor

$$\frac{P_n(x)}{Q_m(x)} - \frac{A_m}{x-a_m} = \frac{P_n(x) - A_m(x-a_1)\dots(x-a_{m-1})}{(x-a_1)(x-a_2)\dots(x-a_m)}.$$

Helyettesítsünk a számlálóba $x = a_m$ -et, akkor az A_m - mivel együtthatója nem nulla - megválasztható úgy, hogy a számláló értéke nulla legyen. A nevező is nulla helyettesítési értéket ad, tehát a tört egyszerűsíthető $(x - a_m)$ -mel, egyszerűsítés után a nevező fokszáma $m - 1$, a számláló fokszáma ennél kisebb lesz, tehát az indukciós feltétel alkalmazható rá.

Mivel az ilyen alakú racionális törtfüggvények parciális törtekre bonthatók, és az egyes törtek integrálása könnyen elvégezhető, az egyetlen kérdés az, hogy a parciális törtekre bontás miként kivitelezhető. A nevező gyökeinek ismeretét feltételezzük. Írjuk fel egyelőre ismeretlen A_1, A_2, \dots, A_m számokkal a

$$\frac{P_n(x)}{Q_m(x)} = \sum_{k=1}^m \frac{A_k}{x-a_k}$$

összefüggést, és szorozzuk meg mindkét oldalt $(x - a_k)$ -val, majd vegyük az $x \rightarrow a_k$ határátmenetet. A jobb oldal folytonos, tehát helyettesíthetünk $x = a_k$ -t, itt minden tag nulla lesz kivéve a k -adikat, ami A_k -t ad. A bal oldal határértékét ki kell számolni, vagy egyszerűsítéssel, vagy L'Hospital szabállyal. Ezt minden k -ra el kell végezni.

13. Impropius integrál

13.1. Az integrálfogalom kiterjesztése

A Riemann-integrál definíciója csak véges intervallumon és csak korlátos függvényekre értelmezhető, mert különben a közelítő összegek végtelenné válnak. Az impropius integrál a Riemann-integrál kiterjesztése a nem korlátos esetekre. (Az impropius latin eredetű szó jelentése "nem valódi".) Az eseteket célszerű külön választani.

D. Az $f: (a, +\infty) \rightarrow \mathbf{R}$ függvény legyen az $(a, +\infty)$ minden véges részintervallumában integrálható, akkor, ha a határérték létezik, akkor létezik az alábbi impropius integrál:

$$\int_a^{\infty} f(x)dx = \lim_{b \rightarrow \infty} \int_a^b f(x)dx.$$

P. Számítsuk ki az $\int_0^{\infty} e^{-x} dx$ impropius integrált.

$$\int_0^{\infty} e^{-x} dx = \lim_{b \rightarrow \infty} \int_0^b e^{-x} dx = \lim_{b \rightarrow \infty} \left[-e^{-x} \right]_0^b = \lim_{b \rightarrow \infty} (1 - e^{-b}) = 1.$$

D. Ha az $f: (a, b) \rightarrow \mathbf{R}$ függvény az (a, b) minden zárt részintervallumán Riemann szerint integrálható (de az (a, b) -n esetleg nem korlátos), akkor, ha a határérték létezik,

$$\int_a^b f(x)dx = \lim_{d \rightarrow b-0} \lim_{c \rightarrow a+0} \int_c^d f(x)dx.$$

Mindkét definícióban foglalt, határértékkal értelmezett integrált közös néven improprius integrálnak nevezzük. Mivel az integrált határértékkal definiáltuk, beszélhetünk az integrál létezésével azonos értelemben az integrál konvergenciájáról is. Ha a második definícióban f korlátos (a, b) -n, akkor ez a Riemann integrál egy folytonossági tulajdonsága, tehát nem mond ellent a Riemann integrál definíciójának. Az integrálás tanult szabályai improprius integrálokra is érvényben maradnak.

P. Számítsuk ki az $\int_0^1 \frac{1}{\sqrt{x}} dx$ improprius integrált.

$$\int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{c \rightarrow +0} \int_c^1 \frac{1}{\sqrt{x}} dx = \lim_{c \rightarrow +0} [2\sqrt{x}]_c^1 = \lim_{c \rightarrow +0} (2 - \sqrt{c}) = 2.$$

13.2. Integrálkritérium sorok konvergenciájára

Az improprius integrálokkal megadható bizonyos típusú sorok konvergenciájának szükséges és elégséges feltétele.

T. Legyen $f(x): [1, \infty) \rightarrow \mathbf{R}$, monoton csökkenő és pozitív függvény. A $\sum_{n=1}^{\infty} f(n)$ akkor és csak

akkor konvergens, ha az $\int_1^{\infty} f(x)dx$ improprius integrál konvergens.

\Rightarrow A sor S_n részletösszege felső közelítő összege az $\int_1^{n+1} f(x)dx$ integrálnak egész

osztópontokat választva, így $\int_1^{n+1} f(x)dx \leq S_n \leq S = \lim S_n$, vagyis az integrálok sorozata

monoton növekedő és korlátos, tehát konvergens.

\Leftarrow Az $\int_1^n f(x)dx$ alsó közelítő összege egész osztópontokra vonatkoztatva $\sum_{k=2}^n f(k)$, tehát

$$\sum_{k=2}^n f(k) \leq \int_1^n f(x)dx \leq \int_1^{\infty} f(x)dx,$$

így a részletösszegek sorozata monoton növekedő és korlátos, tehát konvergens.

P. A $\sum_{n=2}^{\infty} \frac{1}{n \ln n}$ sor divergens. Meg kell nézni az improprius integrált:

$$\int_2^a \frac{1}{x \ln x} dx = [\ln \ln x]_2^a = \ln \ln a - \ln \ln 2 \rightarrow +\infty.$$

14. Az integrálszámítás alkalmazásai

14.1. Területszámítás

Mivel a határozott integrált úgy vezettük be, hogy jelentése a görbe alatti terület, kézenfekvő, hogy az integrálást területszámításra fel tudjuk használni.

D. Ha az $A \subset \mathbf{R}^2$ halmazhoz megadható két, $[a, b]$ intervallumon értelmezett, integrálható függvény - jelölésük legyen f_1 és f_2 , - úgy hogy $f_1(x) \leq f_2(x)$, és az $A = \{(x, y): x \in [a, b], f_1(x) \leq y \leq f_2(x)\}$ alakú, akkor A -t *normáltartomány*-nak nevezzük.

T. Az f_1 és f_2 függvényekkel megadott normáltartomány területe

$$T = \int_a^b f_2(x) - f_1(x) dx.$$

B. T felírható két integrál különbségként, ami átalakítható a fenti alakba.

Általában az $A \subset \mathbf{R}^2$ tetszőleges ponthalmaznak a területe nem értelmezhető. A terület értelmezése a Riemann-integrállal összhangban úgy végezhető el, hogy definiáljuk a külső területet, és a belső területet, és ha a kettő egyenlő, akkor létezik a halmaz területe.

D. Az $A \subset \mathbf{R}^2$ korlátos halmaz *külső területe* a halmazt lefedő véges sok téglalap területösszegének az infimuma. Az $A \subset \mathbf{R}^2$ korlátos halmaz *belső területe* az A részhalmazát képező, véges sok, egymásba nem nyúló téglalapok összterületének a szuprémuma. Az "egymásba nem nyúló téglalapok" azt jelenti, hogy nincs közös belső pontjuk. Az A halmaznak akkor létezik a *területe*, ha a külső területe (ami a korlátosság miatt véges) megegyezik a belső területtel.

14.2. Forgástestek köbtartalma

Az $[a, b]$ intervallumon értelmezett $f(x)$ függvény legyen nemnegatív. Ha a görbét az x -tengely körül megforgatjuk, akkor forgásfelületet, ha a görbe alatti tartományt forgatjuk meg, akkor forgástestet kapunk.

Osszuk fel az $[a, b]$ intervallumot az $x_0 = a < x_1 < x_2 < \dots < x_{n-1} < x_n = b$ számokkal n részre, és jelölje M_i az f függvény szuprémumát, m_i az f infimumát az (x_{i-1}, x_i) intervallumon. A forgástest V köbtartalma a következő két érték közé esik:

$$\sum_{i=1}^n m_i^2 \pi (x_i - x_{i-1}) \leq V \leq \sum_{i=1}^n M_i^2 \pi (x_i - x_{i-1}).$$

A két érték az $\pi \int_a^b f^2(x) dx$ integrál alsó és felső közelítő összege, ha f^2 integrálható függvény, akkor az egyetlen szám, ami mindig a két közelítő összeg közé esik, az integrál értéke, azaz

$$V = \pi \int_a^b f^2(x) dx.$$

15. Kettős integrál

15.1. A kettős integrál definíciója

Az $A \subset \mathbf{R}^2$ ponthalmazról (amit itt tartománynak fogunk nevezni) feltételezzük, hogy létezik a területe. Az A -n értelmezett $f(x, y)$ függvény integrálját az egyváltozós esethez hasonlóan fogjuk értelmezni.

D. Az A tartományt osszuk fel véges sok részre, legyenek ezek a részek A_1, A_2, \dots, A_n . Tételezzük fel, hogy A_1, A_2, \dots, A_n területe létezik, és mindegyik átmérője kisebb δ -nál (azaz A_i bármely két pontjának a távolsága kisebb δ -nál). Ezt nevezzük A δ -nál finomabb felosztásának. Legyen A_i -n $f(x, y)$ szuprémuma M_i , infimuma m_i és jelöljük A_i területét $t(A_i)$ -vel, akkor az

$$\iint_A f(x, y) dx dy$$

integrál alsó közelítő összege $\sum_{i=1}^n m_i t(A_i)$, míg a felső közelítő összege $\sum_{i=1}^n M_i t(A_i)$. Ha

$\delta \rightarrow 0$ esetén a felső összeg és az alsó összeg közös határértékhez tart, akkor $f(x, y)$ az A -n integrálható és integrálja a közös határérték.

Az $\iint_A f(x, y) dx dy$ integrál jelentése nyilván az $f(x, y)$ felület alatti, de A fölé eső térrész térfogata.

15.2. A kettős integrál kiszámítása

Az alábbi tétel szerint, elég általános esetben, a kettős integrál átalakítható kétszeres integrálássá, ami a tanult technikával többnyire megoldható. A tételt először arra az esetre mondjuk ki, amikor A téglalap.

T. Legyen $A = \{(x, y): a < x < b, c < y < d\}$, és $f(x, y)$ az A -n integrálható függvény. Tegyük fel továbbá, hogy minden rögzített $y \in (c, d)$ -re a

$$g(y) = \int_a^b f(x, y) dx$$

integrál létezik, akkor $g(y)$ is integrálható (c, d) -n, és

$$\iint_A f(x, y) dx dy = \int_c^d g(y) dy.$$

B. Osszuk fel mindkét intervallumot az $x_0 = a < x_1 < x_2 < \dots < x_{m-1} < x_m = b$, ill. $y_0 = c < y_1 < y_2 < \dots < y_{n-1} < y_n = d$ osztópontokkal. Az $(x_{i-1}, x_i]$ és az $(y_{j-1}, y_j]$ intervallumok által meghatározott téglalapban legyen f szuprémuma M_{ij} , infimuma m_{ij} , akkor felírhatjuk, hogy a $g(y)$ -t előállító integrál minden $y_j^* \in (y_{j-1}, y_j]$ esetén

$$\sum_{i=1}^m m_{ij}(x_i - x_{i-1}) \leq g(y_j^*) = \int_a^b f(x, y_j^*) dx \leq \sum_{i=1}^m M_{ij}(x_i - x_{i-1}),$$

mert az integrál felső közelítő összegében szereplő $\sup_{x \in (x_{i-1}, x_i]} f(x, y) \leq M_{ij}$, és hasonló állítás

igaz az alsó közelítő összegre is. Szorozzuk meg mindkét oldalt $(y_j - y_{j-1})$ -gyel és összegezzük j -re, akkor

$$\sum_{j=1}^n \sum_{i=1}^m m_{ij}(x_i - x_{i-1})(y_j - y_{j-1}) \leq \sum_{j=1}^n g(y_j^*)(y_j - y_{j-1}) \leq \sum_{j=1}^n \sum_{i=1}^m M_{ij}(x_i - x_{i-1})(y_j - y_{j-1})$$

A bal oldal a kettős integrál alsó, a jobb oldal a felső közelítő összege, jelöljük ezeket s -sel és

S -sel, míg az $\int_c^d g(y) dy$ mindkét közelítő összege is - az előző képletsor szerint - s és S közé

esik. Mivel s és S egyaránt a kettős integrál értékéhez tart, az $\int_c^d g(y) dy$ közelítő összegei is

ezt teszik, tehát a két integrál egyenlő.

A tétel állítása a következő formában használatos:

$$\iint_A f(x, y) dx dy = \int_c^d \left(\int_a^b f(x, y) dx \right) dy,$$

ahol a zárójel természetesen mellőzhető. Az integrálások sorrendje - ha a tétel feltételei teljesülnek - felcserélhető.

Legyen most A a $\varphi_1(x)$ és a $\varphi_2(x)$ által meghatározott, az (a, b) intervallumhoz tartozó normáltartomány. Az $f(x, y)$ legyen az A -n integrálható. Fedjük le A -t a koordináta tengelyekkel párhuzamos oldalú N téglalappal és vezessük be azt a $g(x, y)$ függvényt, amely A -n megegyezik f -fel, A -n kívül pedig 0, akkor

$$\iint_A f dx dy = \iint_A g dx dy = \iint_N g dx dy = \int_a^b \int_c^d g(x, y) dy dx = \int_{\varphi_1(x)}^{\varphi_2(x)} \int_a^b f(x, y) dy dx.$$

Ez a formula tekinthető a normáltartományokra vonatkozó kettős integrálok kiszámítási formulájának.

16. Az n -dimenziós euklideszi tér

16.1. Vektorműveletek

Az n -dimenziós euklideszi tér pontjait az (x_1, x_2, \dots, x_n) valós számokból álló rendezett szám n -esek alkotják. A tér pontjait vektoroknak is nevezzük és $\mathbf{x} = (x_1, x_2, \dots, x_n)$ formában jelöljük. n -et a tér dimenziójának nevezzük. Az n -dimenziós euklideszi tér jelölése \mathbf{R}^n .

A vektorok között műveleteket definiálhatunk. Két *azonos dimenziójú* vektor összeadható és bármely vektor skalárral (számmal) megszorozható, ha $\mathbf{a} = (a_1, a_2, \dots, a_n)$ és $\mathbf{b} = (b_1, b_2, \dots, b_n)$,

akkor $\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$ és $c\mathbf{a} = (ca_1, ca_2, \dots, ca_n)$. A vektorok kivonásának a művelete visszavezethető a (-1)-gyel való szorzás és az összeadás műveletére. Mivel a műveletek a koordinátánkénti összeadást ill. szorzást jelentik, a vektorok összeadása kommutatív és asszociatív művelet, a skalárral való szorzásra nézve pedig disztributív.

D. Ha valmely halmazban a fenti tulajdonságú műveletek elvégezhetők, a halmazt *lineáris térnek* vagy *vektortérnek* nevezzük.

Fontos, vektorok közötti művelet az un. skalárszorzás. Az elnevezés arra utal, hogy a szorzás eredménye szám, azaz skalár mennyiség.

D. Az \mathbf{a} és \mathbf{b} vektorok skalárszorzata:

$$\mathbf{a}\mathbf{b} = a_1b_1 + a_2b_2 + \dots + a_nb_n.$$

Ha egy árukészletben az egyes áruból x_1, x_2, \dots, x_n mennyiségünk van, más szóval a készletvektor $\mathbf{x} = (x_1, x_2, \dots, x_n)$, és a hozzátartozó árak a_1, a_2, \dots, a_n , vagy másképpen fogalmazva az árvektor $\mathbf{a} = (a_1, a_2, \dots, a_n)$, akkor a készlet értéke $\mathbf{a}\mathbf{x}$. Ebből is látható a fogalom rendkívüli fontossága a közgazdaságban.

Könnyen igazolhatóan a skalárszorzás kommutatív, a skalárral való szorzással asszociatív, az összeadással disztributív tulajdonságokkal rendelkező művelet. Képletben:

$$\begin{aligned} \mathbf{a}\mathbf{b} &= \mathbf{b}\mathbf{a}, \\ c(\mathbf{a}\mathbf{b}) &= (c\mathbf{a})\mathbf{b}, \\ c(\mathbf{a} + \mathbf{b}) &= c\mathbf{a} + c\mathbf{b}, \\ (c + d)\mathbf{a} &= c\mathbf{a} + d\mathbf{a}, \\ c(\mathbf{a} + \mathbf{b}) &= c\mathbf{a} + c\mathbf{b}. \end{aligned}$$

M. Két vektor vektoriális szorzata csak háromdimenziós térben definiálható, ezzel most nem foglalkozunk.

16.2. Az n -dimenziós tér geometriája

D. Az \mathbf{a} vektor normája - jelölésben $\|\mathbf{a}\|$ - az

$$\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$$

kifejezéssel értelmezhető.

Az $\|\mathbf{a}\|$ $n = 1, 2$ és 3 esetén természetesen az \mathbf{a} vektor hossza, így a norma a vektor hosszának az általánosítása, ezért a norma szó helyett a hosszúságot is használhatjuk.

Az $\|\mathbf{a}\|$ kifejezhető a skalárszorzattal: $\|\mathbf{a}\|^2 = \mathbf{a}\mathbf{a}$.

T (Cauchy-Schwarz egyenlőtlenség). Tetszőleges \mathbf{a} és \mathbf{b} vektorokra $|\mathbf{a}\mathbf{b}| \leq \|\mathbf{a}\| \cdot \|\mathbf{b}\|$.

B. Számítsuk ki az $\|\mathbf{a} + \lambda\mathbf{b}\|^2$ kifejezést és használjuk fel, hogy a kifejezés egyetlen λ értékre sem lehet negatív.

$$\|\mathbf{a} + \lambda\mathbf{b}\|^2 = (\mathbf{a} + \lambda\mathbf{b})(\mathbf{a} + \lambda\mathbf{b}) = \|\mathbf{a}\|^2 + 2\lambda\mathbf{a}\mathbf{b} + \lambda^2\|\mathbf{b}\|^2 \geq 0,$$

és egy másodfokú függvény (ami most λ -nak a függvénye) csak úgy lehet minden λ -ra nemnegatív, ha a diszkriminánsa negatív, vagy nulla, azaz

$$4(\mathbf{a}\mathbf{b})^2 - 4\|\mathbf{a}\|^2 \cdot \|\mathbf{b}\|^2 \leq 0,$$

ami az állítással ekvivalens.

T (Háromszög egyenlőtlenség). $\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$.

B. A Cauchy-Schwarz egyenlőtlenséget felhasználva

$\|\mathbf{a} + \mathbf{b}\|^2 = (\mathbf{a} + \mathbf{b})(\mathbf{a} + \mathbf{b}) = \|\mathbf{a}\|^2 + 2\mathbf{a}\mathbf{b} + \|\mathbf{b}\|^2 \leq \|\mathbf{a}\|^2 + 2\|\mathbf{a}\| \cdot \|\mathbf{b}\| + \|\mathbf{b}\|^2 = (\|\mathbf{a}\| + \|\mathbf{b}\|)^2$,
ami az állítással ekvivalens.

Ha \mathbf{a} és \mathbf{b} egy háromszög két oldalvektora, akkor $\mathbf{a} - \mathbf{b}$ a harmadik oldal vektora. Számítsuk ki az előbbiekhöz hasonlóan $\|\mathbf{a} - \mathbf{b}\|^2$ -t:

$$\|\mathbf{c}\|^2 = \|\mathbf{a} - \mathbf{b}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\mathbf{a}\mathbf{b},$$

amit a háromszögre vonatkozó koszinusz-tétellel összehasonlítva ($n = 2$ és 3 esetén)

$$\cos \gamma = \frac{\mathbf{a}\mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|},$$

így ezzel a képlettel magasabb dimenzió esetén is értelmezhető és kiszámítható két vektor szöge. A képletből látható, hogy $\gamma = 90^\circ$ akkor és csak akkor, ha $\mathbf{a}\mathbf{b} = 0$.

D. Két vektor, \mathbf{a} és \mathbf{b} , ortogonális (merőleges), ha $\mathbf{a}\mathbf{b} = 0$.

D. \mathbf{R}^n alterének nevezzük bármely olyan nem üres részhalmazát, amely maga is lineáris tér.

P. Például \mathbf{R}^3 altere minden olyan sík, amely az origón áthalad (kétdimenziós altér), vagy minden egyenes, amely az origón áthalad (egydimenziós altér). Az origót elkerülő sík nem altér, mert $\mathbf{0}$ mindig eleme kell legyen az alternek.

16.3. Lineáris függetlenség

D. A $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ vektorok *lineárisan függetlenek*, ha a $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k = \mathbf{0}$ egyenlőség csak úgy teljesülhet, ha $c_1 = c_2 = \dots = c_k = 0$ ($\mathbf{0}$ az előző képletben a nullvektort jelöli, melynek minden komponense 0).

Ha a $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ vektorokra a lineáris függetlenség nem teljesül, azaz lineárisan nem függetlenek, akkor van olyan i , melyre $c_i \neq 0$, és ekkor

$$\mathbf{v}_i = -\frac{c_1}{c_i}\mathbf{v}_1 - \frac{c_2}{c_i}\mathbf{v}_2 - \dots - \frac{c_k}{c_i}\mathbf{v}_k,$$

vagyis valamelyik vektor a többi lineáris kombinációjával előállítható, más szóval valamilyen i -re a \mathbf{v}_i benne van a többi vektor által generált altérben.

M. Fontos tulajdonsága a lineárisan független $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ vektoroknak, hogy ha valamely \mathbf{a} előállítható $\mathbf{a} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k$ alakban, akkor ez az előállítás egyértelmű. Ha ugyanis $\mathbf{a} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \dots + d_k\mathbf{v}_k$ lenne, akkor $(c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + \dots + (c_k - d_k)\mathbf{v}_k = \mathbf{0}$ a lineáris függetlenség miatt csak úgy teljesülhetne, hogy minden i -re $c_i = d_i$.

D. Egy altér dimenziója az altérben lévő lineárisan független vektorok maximális száma.

17. Mátrixok

17.1. Műveletek mátrixokkal

D. A téglalap formába rendezett, valós számokból álló halmazokat *mátrixnak* nevezzük. Általános alakja:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{bmatrix}.$$

A megadott mátrixnak n sora és m oszlopa van, röviden $n \times m$ -es mátrixnak nevezzük. Az egyes soraiból képezett vektorokat *sorvektoroknak*, az oszlopokból képzetteket *oszlopvektoroknak* nevezzük. Az a_{ij} jelölésnél az első index mindig a sorindex, a második az oszlopindex. Ha $m \geq n$, akkor az $a_{11}, a_{22}, a_{33}, \dots, a_{mm}$ ($m \leq n$ esetén az $a_{11}, a_{22}, a_{33}, \dots, a_{mm}$) elemek a mátrix *főátlóját* alkotják. A mátrixokat vastagított nagy betűvel jelöljük.

D. Két mátrix *egyenlő*, ha méretük megegyezik (vagyis mindkettő $n \times m$ -es mátrix), és az azonos helyzetű elemek egyenlők a két mátrixban.

D. Mátrixok *szorzása számmal* úgy történik, hogy minden elemét megszorozzuk. *Összeadni* két mátrixot csak akkor lehet, ha méretük megegyezik (vagyis mindkettő $n \times m$ -es mátrix), ekkor a megfelelő elemeket kell összeadni.

Mivel mindkét mátrix-művelet az egyes elemeken történő szokásos értelmű művelet, nyilván kommutatív, asszociatív és disztributív: $(ab)\mathbf{A} = a(b\mathbf{A})$, $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$, $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$, $(a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$, $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$.

A mátrixok szorzása ennél jóval bonyolultabb, gyakorlást igénylő művelet.

D. Két mátrixot, \mathbf{A} -t és \mathbf{B} -t *összeszorozni* csak akkor lehet, ha az első mátrix oszlopmérete megegyezik a második sorméretével. Az \mathbf{AB} mátrix i -edik sorának j -edik eleme a \mathbf{A} mátrix i -edik sorának és a \mathbf{B} -mátrix j -edik oszlopának a skalárszorzata. (Jegyezzük meg: a mátrix-szorzás *sor - oszlop szorzás!*) Ebből következik, hogy egy $n \times p$ -s és egy $p \times m$ -es mátrix szorzata $n \times m$ -es mátrix lesz.

A mátrix-szorzás nem kommutatív, hiszen általában nincs is értelmezve a fordított sorrendű szorzás, mert a méretre vonatkozó kikötést nem fogja teljesíteni. Ha négyzetes ($n \times n$ -es) mátrixokra nézzük, akkor értelmes lesz, de általában $\mathbf{AB} \neq \mathbf{BA}$. Ennek ellenére a skalárral való szorzásra nézve asszociatív, az összeadásra nézve disztributív: $a(\mathbf{AB}) = (a\mathbf{A})\mathbf{B}$, $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ és $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$. (A disztributív szabály a vektorokra vonatkozó szabály következménye.) A mátrix szorzás asszociatív, de részletes bizonyításra szorul.

T. Ha az \mathbf{A} , \mathbf{B} , \mathbf{C} mátrixok méretei az \mathbf{AB} és a \mathbf{BC} szorzások elvégzését lehetővé teszik, akkor $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

B. Jelöljük az \mathbf{A} mátrix elemeit a_{ij} -vel, a \mathbf{B} mátrix elemeit b_{ij} -vel és a \mathbf{C} mátrixét c_{ij} -vel, a mátrixok méreteit nem tüntetjük fel az összegzésekben. Az \mathbf{AB} mátrix elemei legyenek α_{ij} , a \mathbf{BC} mátrix elemei β_{ij} , akkor

$$\alpha_{ij} = \sum_k a_{ik} b_{kj}$$

és

$$\beta_{kl} = \sum_j b_{kj} c_{jl}.$$

Az $(\mathbf{AB})\mathbf{C}$ mátrix i -edik sorának l -edik eleme:

$$\begin{aligned} \sum_j \alpha_{ij} c_{jl} &= \sum_j \left(\sum_k a_{ik} b_{kj} \right) c_{jl} = \sum_j \sum_k a_{ik} b_{kj} c_{jl} = \\ &= \sum_k \sum_j a_{ik} b_{kj} c_{jl} = \sum_k a_{ik} \left(\sum_j b_{kj} c_{jl} \right) = \sum_k a_{ik} \beta_{kl}, \end{aligned}$$

ami az $\mathbf{A}(\mathbf{BC})$ mátrix i -edik sorának l -edik eleme.

D. Mátrix transzponáltja a sorok és oszlopok felcserélésével (vagy másképpen, a főátlóra való tükrözéssel) kapott mátrix. \mathbf{A} transzponáltjának a jelölése \mathbf{A}^* .

T. $(\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$, és $(\mathbf{AB})^* = \mathbf{B}^* \mathbf{A}^*$.

B. Az első állítás a transzponálás elvégzésével közvetlenül látható. A másodikonál arra kell gondolni, hogy a mátrix-szorzás sor-oszlop szorzás, de a transzponálás felcseréli a sorokat és az oszlopokat. Az $(\mathbf{AB})^*$ mátrix i -edik sorának j -edik eleme megegyezik az \mathbf{AB} mátrix j -edik sorának i -edik elemével, amit úgy kapok meg, hogy a \mathbf{B} i -edik oszlopát skalárisan szorzom \mathbf{A} j -edik sorával, vagyis \mathbf{B}^* i -edik sorát szorzom \mathbf{A}^* j -edik oszlopával, ami $\mathbf{B}^* \mathbf{A}^*$ i -edik sorának j -edik eleme.

17.2. Lineáris transzformáció

A mátrix műveleteknél a vektorokat célszerű oszlopvektoroknak tekinteni, azaz $n \times 1$ -es mátrixnak. Ha mégis sorvektornak vennénk, akkor azt transzponálással jelezzük. Legyen $\mathbf{A} = (a_{ij})$ $n \times m$ -es mátrix, és

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

m -dimenziós oszlopvektor, akkor az $\mathbf{y} = \mathbf{Ax}$ kifejezés mátrixok szorzásával kiszámítható és a szorzás eredménye az \mathbf{y} -nal jelölt n -dimenziós oszlopvektor. Ezt úgy is kifejezhetjük, hogy az $\mathbf{y} = \mathbf{Ax}$ függvénykapcsolatot létesít és leképezi \mathbf{R}^m -et \mathbf{R}^n -be. A leképezés lineáris, hiszen $\mathbf{x}_1 + \mathbf{x}_2$ -hez $\mathbf{Ax}_1 + \mathbf{Ax}_2$ -t, $c\mathbf{x}$ -hez $c\mathbf{Ax}$ -et rendeli hozzá.

Írjuk ki részletesen a hozzárendelés módját:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1m}x_m \\ y_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2m}x_m \\ &\dots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nm}x_m. \end{aligned}$$

Ebből látható például, hogy a eredeti tér koordináta egységvektorait, a $(0, 0, \dots, 0, 1, 0, \dots, 0)$ alakú vektorokat, melyeknek csak az i -edik koordinátája 1, a többi 0, az i -edik oszlopvektorba viszi át a leképezés.

17.3. Mátrix rangja

D. A mátrix rangja a lineárisan független oszlopvektorok száma (pontosabban az oszlopvektorok közül kiválasztható lineárisan független vektorok maximális száma). Az A mátrix rangját $\text{rang}(A)$ -val jelöljük.

T. Ha az $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_m$ vektorok között pontosan k darab lineárisan független vektor van, akkor ugyanennyi lineárisan független vektor van ($c \neq 0$ -t feltételezve) a

$$c\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_m$$

és az

$$\mathbf{a}_1 + \mathbf{a}_j, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_m \quad (1 \leq j \leq m)$$

vektorok között.

B. Elég azt bizonyítani, hogy a lineárisan független vektorok száma nem csökken, ugyanis, ha nőtt volna, akkor ugyanilyen lépéssel az eredeti vektorrendszer helyreállítható, és ennél a lépésnél a lineárisan független vektorok száma csökkenne.

Válasszunk ki k darab lineárisan független vektort, feltehető, hogy ezek $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_k$, ugyanis, ha \mathbf{a}_1 nem lenne köztük, akkor ezek változatlanul szerepelnének a módosított vektorok között is, tehát legalább k lineárisan független vektor van köztük.

Az első esetben könnyen látható, hogy $c\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_k$ lineárisan függetlenek, ugyanis vizsgáljuk meg a $c_1c\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 + \dots + c_k\mathbf{a}_k = \mathbf{0}$ egyenlőséget. Mivel $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_k$ lineárisan függetlenek, $c_1c = 0, c_2 = 0, \dots, c_k = 0$, amiből látszik, hogy az egyenlőség csak úgy teljesülhet, ha valamennyi $c_i = 0$, vagyis $c\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_k$ lineárisan függetlenek.

A második esetben azt fogjuk bizonyítani, hogy vagy az $\mathbf{a}_1 + \mathbf{a}_j, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_k$ vektorok, vagy az $\mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_k, \mathbf{a}_j$ vektorok lineárisan függetlenek. Tegyük fel az ellenkezőjét, vagyis azt, hogy mindkét vektorrendszer lineárisan összefüggő. Így léteznek olyan c_i számok (nem minden $c_i = 0$ kikötéssel), hogy

$$c_1(\mathbf{a}_1 + \mathbf{a}_j) + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 + \dots + c_k\mathbf{a}_k = \mathbf{0},$$

és $c_1 \neq 0$, mert az $\mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_k$ vektorok lineárisan függetlenek. \mathbf{a}_j tehát kifejezhető:

$$\mathbf{a}_j = -\frac{1}{c_1}(c_2\mathbf{a}_2 + \dots + c_k\mathbf{a}_k) - \mathbf{a}_1.$$

Hasonlóan, ha $\mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_k, \mathbf{a}_j$ vektorok lineárisan összefüggők, akkor \mathbf{a}_j kifejezhető

$$\mathbf{a}_j = \sum_{i=2}^k d_i\mathbf{a}_i = -\frac{1}{c_1}(c_2\mathbf{a}_2 + \dots + c_k\mathbf{a}_k) - \mathbf{a}_1,$$

és az összevetésből az látszik, hogy az $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_k$ vektorok lineárisan összefüggők, hiszen pl. \mathbf{a}_1 együtthatója biztosan nem zérus, és ez ellentmondás.

A tétel nagyon fontos eljárást ad mátrixok rangjának a kiszámításához. Ha egy mátrix valamely oszlopának c -szeresét hozzáadjuk egy másik oszlophoz, akkor a rangja nem változik. Ugyanez az eljárás sorokra is elvégezhető, mint a következő tételben megmutatjuk. Az eljárással - amit elemi sorműveletnek nevezünk - sorra gyárthatók a nullák a mátrixban a rang megváltozása nélkül, és végül a rang leolvasható.

T. Legyen $\mathbf{a}_i = (a_{i1}, a_{i2}, a_{i3}, \dots, a_{im})$ és $\mathbf{b}_i = (a_{i1} + ca_{i2}, a_{i2}, a_{i3}, \dots, a_{im})$ ($i = 1, 2, \dots, m$). Az $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_m$ vektorok akkor és csak akkor lineárisan függetlenek, ha a $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \dots, \mathbf{b}_m$ vektorok is azok.

B. Lineáris függőségre bizonyítjuk az állítást. Legyenek az $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_m$ vektorok lineárisan összefüggők, akkor $\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \lambda_3\mathbf{a}_3 + \dots + \lambda_m\mathbf{a}_m = \mathbf{0}$ úgy, hogy nem minden együttható 0. Ez az összefüggés minden koordinátában igaz, vagyis

$$\lambda_1a_{1i} + \lambda_2a_{2i} + \lambda_3a_{3i} + \dots + \lambda_ma_{mi} = 0$$

minden i -re, ebből következik, hogy $\lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2 + \lambda_3 \mathbf{b}_3 + \dots + \lambda_m \mathbf{b}_m = \mathbf{0}$ minden koordinátájában teljesül, vagyis $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \dots, \mathbf{b}_m$ is lineárisan összefüggő.
A fordított állításhoz $-c$ -vel alkalmazzuk a bizonyított részt.

Nézzük meg egy példán keresztül. Számítsuk ki az A mátrix rangját! (Az egyes lépések magyarázata a képletsor után található. Az egyes átalakításokat \sim jellel kötjük össze, ami itt csak azt jelenti, hogy a két mátrix rangja egyenlő.)

$$\begin{aligned}
 A = \begin{bmatrix} 4 & 2 & 1 & 3 \\ 0 & -1 & -2 & 2 \\ -1 & 1 & 0 & -2 \\ 2 & 4 & 1 & -1 \end{bmatrix} &\sim \begin{bmatrix} 4 & 2 & -3 & 7 \\ 0 & -1 & 0 & 0 \\ -1 & 1 & -2 & 0 \\ 2 & 4 & -7 & 7 \end{bmatrix} \sim \begin{bmatrix} 4 & 0 & -3 & 7 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & -2 & 0 \\ 2 & 0 & -7 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 4 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & -2 & 0 \\ 2 & 0 & -7 & 7 \end{bmatrix} \sim \\
 &\sim \begin{bmatrix} 2 & 0 & 4 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & -2 & 0 \\ 0 & 0 & 0 & 7 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & -2 & 0 \\ 0 & 0 & 0 & 7 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 \end{bmatrix}
 \end{aligned}$$

1. A második oszlop (-2) -szeresét adjuk hozzá a harmadikhoz és a kétszeresét adjuk hozzá a negyedikhez.
2. Mivel a második sorban minden elem - egy kivételével - nulla, ennek az elemnek az oszlopában a többi elem nullázható (hozzáadva a második sor alkalmas többszörösét az első, majd a harmadik és a negyedik sorhoz). Minden más elem megmarad.
3. Adjuk hozzá a negyedik sor (-1) -szeresét az első sorhoz.
4. A negyedik sor többi eleme nullázható.
5. Adjuk hozzá a harmadik sor kétszeresét az elsőhöz.
6. Nullázzuk a harmadik sort.
7. Itt már minden sorban és oszlopban legfeljebb egy nem nulla elem van - az eljárás véget ért. A mátrixban minden olyan oszlop, amelyben van nem nulla elem, nyilván lineárisan független egymástól, míg a 0 elemekből álló oszlopvektorok függenek a többitől. A mátrix rangja tehát a nem nulla elemek száma, azaz 3.

Mivel ez az eljárás mindig elvégezhető néhány általános következtetést levonhatunk.

1. Minden mátrixban ugyanannyi a lineárisan független sorvektorok száma, mint a lineárisan független oszlopvektorok száma. Így a transzponált mátrix rangja egyenlő az eredeti rangjával: $\text{rang}(A) = \text{rang}(A^*)$.
2. Az $n \times m$ -es A mátrix rangja nem lehet több, mint $\min(n, m)$.
3. Ha a végállásban kapott csupa nulla oszlopok helyén álló oszlopokkal oszlopműveletet nem végeztünk, akkor a nem nulla oszlopoknak megfelelő oszlopok az eredeti mátrixban lineárisan függetlenek. Hasonló állítás igaz a sorokra is.

17.4. Lineáris egyenletrendszer megoldhatósága

Ha visszalapozunk a lineáris transzformáció részletesen felírt alakjára, akkor világos, hogy egy n egyenletből álló m ismeretlenes lineáris egyenletrendszer általános alakja

$$A\mathbf{x} = \mathbf{b},$$

ahol, A $n \times m$ -es mátrix, $\mathbf{x}^* = (x_1, x_2, \dots, x_m)$ az ismeretlenekből képezett vektor (\mathbf{x}^* itt az \mathbf{x} oszlopvektor transzponáltját jelöli, helykímélés céljából célszerű így írni) és $\mathbf{b}^* = (b_1, b_2, \dots, b_m)$ az állandó tagokból képezett vektor.

T. Az $\mathbf{Ax} = \mathbf{b}$ lineáris egyenletrendszer akkor és csak akkor oldható meg, ha $\text{rang}([\mathbf{A}, \mathbf{b}]) = \text{rang}(\mathbf{A})$. ($[\mathbf{A}, \mathbf{b}]$ itt a \mathbf{b} oszlopvektorral kibővített \mathbf{A} mátrixot jelenti.)

B. Jelöljük az \mathbf{A} mátrix oszlopvektorait $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ -mel, és az oszlopvektorokból összetett mátrixra használjuk az $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ jelölést.

\Rightarrow Ha az egyenletrendszer megoldható, akkor vannak olyan x_1, x_2, \dots, x_m számok, hogy

$$\mathbf{b} = \sum_{i=1}^m \mathbf{a}_i x_i . \text{ A 17.3 fejezet első tétele alapján } \text{rang}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m) = \text{rang}(\mathbf{a}_1 x_1, \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m) =$$

$\text{rang}(\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2, \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m) = \dots = \text{rang}(\mathbf{b}, \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$, és ezt akartuk bizonyítani.

\Leftarrow Ha $\text{rang}([\mathbf{A}, \mathbf{b}]) = \text{rang}(\mathbf{A}) = k$, akkor kiválasztható \mathbf{A} -ból k darab lineárisan független oszlopvektor, legyen ez pl. $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$, és $[\mathbf{A}, \mathbf{b}]$ -ből $k + 1$ már nem választható ki, vagyis $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m, \mathbf{b}$ már lineárisan összefüggők. Ez azt jelenti, hogy $c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_k \mathbf{a}_k + c_{k+1} \mathbf{b} = \mathbf{0}$, és nem valamennyi $c_i = 0$. A $c_{k+1} = 0$ nem lehet, mert akkor az $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ vektorok lineárisan összefüggők lennének. Ebből

$$\mathbf{b} = -\frac{c_1}{c_{k+1}} \mathbf{a}_1 - \frac{c_2}{c_{k+1}} \mathbf{a}_2 - \dots - \frac{c_k}{c_{k+1}} \mathbf{a}_k ,$$

ami azt jelenti, hogy előállítottuk az egyenletrendszer egy megoldását.

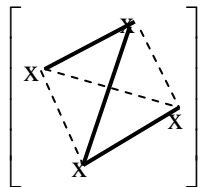
18. Determinánsok

A determináns a négyzetes ($n \times n$ -es) mátrixokhoz rendelt számérték, ezért ebben a fejezetben minden mátrixról feltételezzük, hogy négyzetes mátrix.

18.1. A determináns definíciója

A determináns definíciója meglehetősen bonyolult, nem fogjuk képletszerűen felírni. Később majd adunk képletet is a kiszámítására.

D. Az $n \times n$ -es \mathbf{A} mátrixból válasszunk ki n elemet úgy, hogy egy sorból és egy oszlopból csak egy elem kerüljön kiválasztásra. A kiválasztott elemeket szorozzuk össze, és az összes lehetséges kiválasztásra ezeket a szorzatokat megfelelő előjellel ellátva adjuk össze, így kapjuk meg az \mathbf{A} determinánsának az értékét. Ezeket a szorzatokat (előjelezését nem beleértve) *elemi szorzatoknak* fogjuk nevezni. Az előjelszabály megállapításához rajzoljuk be a mátrixban a kiválasztott elemek helyét:



kössünk össze minden kiválasztott elemet minden másikkal, ha a jobbra felfelé haladó összekötések (az ábrán vastagítva) száma páros, akkor a kiválasztott elemekhez tartozó előjel pozitív, ha páratlan, akkor negatív. \mathbf{A} determinánsának jelölése $|\mathbf{A}|$

Nézzük meg kisebb mátrixokon, hogy mit jelent a definíció.

Ha 2×2 -es mátrix, akkor

$$|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21},$$

vagyis a főátlóbeli elemek szorzata mínusz a mellékátlóbeli elemek szorzata.

Ha 3×3 -as mátrix, akkor

$$|A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{13} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33},$$

ennek a formulának a megjegyzéséhez az ún. *Sarrus-szabály* nyújt segítséget (lásd a gyakorlaton).

Általában a determináns közvetlen kiszámítása nem járható út, mert $n \times n$ -es determináns kiszámításához $n!$ tagot kell előjelezni és kiszámítani. A definícióból mégis közvetlenül következik a determinánsok néhány fontos tulajdonsága.

18.2. A determinánsok elemi tulajdonságai

T1. A determináns értéke nulla, ha egy sora vagy egy oszlopa csupa nullából áll.

B. Minden elemi szorzatban lesz egy 0-s tényező.

T2. A transzponált mátrix determinánusa megegyezik az eredeti determinánusával.

B. Ugyanazokból az elemi szorzatokból tevődik össze és a transzponálással az előjelszabállyal kapott előjel is ugyanaz lesz.

T3. A mátrix két sorát (vagy oszlopát) felcserélve a determináns előjelet vált.

B. Elég szomszédos sorokra bizonyítani, mert nem szomszédos sorok cseréje páratlan sok szomszédos cserével megoldható. A csere után minden elemi szorzatban a jobbra felfelé haladó "összekötések" megmaradnak jobbra felfelé tartónak, kivéve a két felcserélt sor elemeit összekötő vonalat, melynek iránya megváltozik. Így minden elemi szorzat előjelet vált.

T4. A determináns értéke c -szeres lesz, ha a mátrix *egy sorát* (vagy oszlopát) c -vel megszorozzuk.

B. Minden elemi szorzatban egy tényező c -szeresére változik.

T5. A determináns értéke nulla, ha egyik sor (vagy oszlop) a másik c -szerese. Speciálisan, a determináns értéke nulla, ha a mátrix két sora (vagy oszlopa) megegyezik.

B. A második állítás T3 következménye. A második állításból az első már következik T4 felhasználásával.

T6. Ha két mátrix sorai megegyeznek, kivéve az első sort, akkor annak a mátrixnak a determinánusa, amelyet az első sorok elemenkénti összegzéséből és a többi sorok változatlan lemásolásából kapunk, a két eredeti mátrix determinánsának az összege. Ugyanezt elmondhatjuk a mátrix bármely sorára vagy oszlopára vonatkozóan is.

B. Az első sor összeadásával kapott mátrix determinánsának elemi szorzatai beszorzással felbonthatók két tagra, egyik az első, a másik a második determináns egy-egy elemi szorzata. Az előjelszabály mindhárom esetben ugyanazt az előjelet adja.

T7. Ha egy mátrix valamely sorához (oszlopához) a másik sor (ill. oszlop) c -szeresét hozzáadjuk, a determináns értéke nem változik.

B. Az eredményül kapott mátrix a T6-ban szereplő összeadási művelettel két mátrix "összegére" bontható, a determinánsok összeadódnak, de az egyiknek a determinánusa nulla. (Ez az "összeadás" különbözik a mátrixok szokásos összeadási műveletétől!).

T8. Az $n \times n$ -es determináns értéke akkor és csak akkor nulla, ha rangja n -nél kisebb. A determináns értéke ugyanúgy számolható ki, ahogy a mátrixok rangját számoltuk: ha eljutottunk oda, hogy minden sorban és minden oszlopban csak egy nullától különböző elem van, akkor az elemi szorzatok egy kivételével nullák, ez pedig könnyen kiszámítható.

B. Ha a rang n -nél kisebb, akkor az átalakításnál a mátrixnak lesz csupa nullából álló sora, vagyis a determinánsa nulla lesz.

18.3. Kifejtési tétel

D. Az $A = (a_{ij})$ mátrix adott a_{ij} eleméhez tartozó *előjeles al-determinánsa* az i -edik sor és a j -edik oszlop elhagyásával kapott mátrix determinánsa szorozva $(-1)^{i+j}$ -nel. Az a_{ij} elemhez tartozó előjeles al-determinánst a továbbiakban A_{ij} -vel fogjuk jelölni.

A fenti előjelszabályt röviden sakktáblaszabálynak szoktuk említeni, mivel a mátrix megfelelő helyeire a különböző előjeleket beírva a + és a - előjelek úgy váltakoznak, mint a sakktáblán a fekete és a fehér mezők.

T (kifejtési tétel). Az $|A|$ determináns értékét az i -edik sor szerint *kifejtve* a

$$|A| = \sum_{j=1}^n a_{ij} A_{ij}$$

képlettel tudjuk visszavezetni eggyel alacsonyabbrendű determináns kiszámítására. Hasonló tétel érvényes az oszlop szerinti kifejtésre is.

B. Elég az első sor szerinti kifejtésre bizonyítani, hiszen sorcserékkel az i -edik sor felhozható az első sorba, minden sorcserénél a determináns értéke előjelet vált, de ezt a sakktáblaszabály figyelembe veszi, mert az is biztosítja a megfelelő előjel váltást. Az első sor szerinti kifejtés képlete:

$$|A| = \sum_{j=1}^n a_{1j} A_{1j}$$

Az $|A|$ determináns elemi szorzatok előjelezett összege. Minden elemi szorzat tartalmaz első sorbeli elemet. Gyűjtsük össze azokat a tagokat, amelyek a_{11} -et, majd amelyek a_{12} -t, ... tartalmazznak és emeljük ki belőlük az első sorbeli elemet, akkor az a_{1j} szorzója az a_{1j} -hez tartozó al-determináns elemi szorzataiból áll, csupán az előjelszabályt kell ellenőrizni. Az A mátrixból válasszunk ki egy elemi szorzatot, és készítsük el hozzá az előjelszabály eldöntéséhez használt, összekötéseket tartalmazó ábrát. Ha ez az elemi szorzat az a_{1j} elemet tartalmazza az első sorból, akkor $j - 1$ jobbra felfelé haladó összekötés fut be a_{1j} -be. A többi jobbra felfelé haladó összekötés már az al-determinánsban adja meg az elemi szorzat előjelszabályát. Így a két előjel $(-1)^{j-1} = (-1)^{j+1}$ szorzóban különbözik, ami megfelel a sakktáblaszabálynak.

19. Lineáris egyenletrendszerek megoldása

19.1. Inverz mátrix

Jelöljük, mint ezt az előző fejezetben tettük, A_{ij} -vel az a_{ij} elemhez tartozó előjeles aldeterminánst, és legyen

$$\mathbf{B} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}.$$

Számoljuk ki az \mathbf{AB}^* mátrixot. Az eredménymátrix i -edik sorának j -edik eleme az \mathbf{A} mátrix i -edik sorvektorának és a \mathbf{B} mátrix j -edik sorvektorának a skalárszorzata. Ha $i = j$, akkor a skalárszorzat

$$\sum_{k=1}^n a_{ik} A_{ik},$$

ami a kifejtési tétel miatt $|\mathbf{A}|$. Ha $i \neq j$, akkor egy olyan mátrix determinánsára írtuk fel a kifejtési tételt, amelynek a j -edik sorát az i -edikkel azonosra változtattuk meg, de az ilyen mátrix determinánsa nulla. Az eredmény tehát:

$$\mathbf{AB}^* = \begin{bmatrix} |\mathbf{A}| & 0 & \cdots & 0 \\ 0 & |\mathbf{A}| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |\mathbf{A}| \end{bmatrix} = |\mathbf{A}| \mathbf{I},$$

ahol \mathbf{I} az egységmátrix, melynek minden főátlóbeli eleme 1, a többi eleme pedig 0. A $\mathbf{B}^* \mathbf{A}$ szorzat is ugyanezt az eredményt adja.

D. Az \mathbf{A} $n \times n$ -es mátrix inverzének nevezzük azt az \mathbf{A}^{-1} mátrixot, melyre $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

T. Ha az $|\mathbf{A}| \neq 0$, akkor az \mathbf{A} inverze létezik és

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \mathbf{B}^*.$$

T. Ha $|\mathbf{A}| = 0$, akkor az \mathbf{A}^{-1} inverzmátrix nem létezik.

B. Tegyük fel, hogy az inverzmátrix létezik, akkor az $\mathbf{Ax} = \mathbf{b}$ lineáris egyenletrendszer minden $\mathbf{b} \in \mathbf{R}^n$ -re megoldható, hiszen jobbról beszorozva \mathbf{A}^{-1} -gyel $\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Az egyenletrendszer azonban csak akkor oldható meg, ha a kibővített mátrix rangja megegyezik \mathbf{A} rangjával (ld, 17.4.). Ha $|\mathbf{A}| = 0$, akkor rangja kisebb n -nél, így található olyan $\mathbf{b} \in \mathbf{R}^n$, mely lineárisan független \mathbf{A} oszlopvektoraitól, és ekkor a kibővített mátrix rangja nagyobb, mint \mathbf{A} rangja, ami ellentmondást jelent.

19.2. Cramer-szabály

Tekintsük az

$$\mathbf{Ax} = \mathbf{b}$$

lineáris egyenletrendszert, ahol \mathbf{A} $n \times n$ -es mátrix, vagyis az egyenletrendszerünk n egyenletből áll és n ismeretlent tartalmaz. Az előző pontban láttuk, hogy $|\mathbf{A}| \neq 0$ esetén az egyenletrendszer egyértelműen megoldható és megoldása $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Írjuk be az inverz mátrix kiszámítási formuláját ebbe a képletbe:

$$\mathbf{x} = \frac{1}{|\mathbf{A}|} \mathbf{B}^* \mathbf{b},$$

és nézzük, hogy a k -edik koordinátára, x_k -ra mit kapunk. A B^*b vektor k -edik eleme a B mátrix k -edik oszlopának és a b -nek a skalárszorzata:

$$\sum_{i=1}^n A_{ik} b_i,$$

ami a kifejtési tétel szerint olyan mátrix determinánsa, melynek minden eleme megegyezik A megfelelő elemével, de a k -edik oszlopvektora b .

T (Cramer-szabály). Legyen A $n \times n$ -es mátrix, és $|A| \neq 0$, akkor az $Ax = b$ egyenlet egyértelműen megoldható, és megoldása

$$x_k = \frac{|A_k|}{|A|} \quad (k = 1, 2, \dots, n),$$

ahol az A_k mátrixot úgy kapjuk, hogy az A k -edik oszlopát b -re cseréljük.

Ha az egyenletrendszer A mátrixa $n \times m$ -es és $m > n$ (több ismeretlen van, mint egyenlet), és az A mátrix rangja n , akkor válasszuk ki n darab lineárisan független oszlopvektort, a többi oszlopot a hozzátartozó ismeretlenekkel szorozva vigyük át a jobb oldalra és olvasszuk bele a b -be. A jobb oldalra átkerülő ismeretleneknek tetszőleges értéket adva az egyenletrendszer a Cramér-szabállyal megoldható, de mivel minden értékadáshoz tartozik egy megoldás, az egyenletrendszernek végtelen sok megoldása lesz. Az általános megoldás úgy kapható, hogy a jobboldali ismeretleneket paramétereknek tekintve oldjuk meg a feladatot. A szabadon választható paraméterek száma $m - n$.

Ha az egyenletrendszer A mátrixa $n \times m$ -es, ahol n és m tetszőleges, és A rangja r , akkor válasszuk ki A -ból egy $r \times r$ -es A_0 részmatrixot, melynek rangja továbbra is r . Ha a megoldhatóság 17.4.-ben kimondott feltétele teljesül, akkor azok az egyenletek, melyek együtthatóit A_0 -ba nem választottuk be, elhagyhatók, a kapott megoldások ezeket automatikusan ki fogják elégíteni. A többi egyenletre az előző eljárás alkalmazható.

20. Differenciálegyenletek

20.1. Példa modell-alkotással

Készítsünk modellt a népesség alakulására adott területen. $N(t)$ jelölje a népesség számát a t időpontban, akkor a legegyszerűbb modell szerint a népesség számának változása kis időintervallumban arányos a meglévő népességgel és az időintervallum hosszával, vagyis

$$N(t+h) - N(t) = cN(t)h.$$

h -val elosztva az egyenlet mindkét oldalát

$$\frac{N(t+h) - N(t)}{h} = cN(t),$$

majd $h \rightarrow 0$ határátmenetet véve az

$$N'(t) = cN(t)$$

egyenletet kapjuk. A megoldandó feladat az, hogy meg kell határozni azt az $N(t)$ függvényt, amelyik ennek az egyenletnek eleget tesz. Ezt az egyenletet differenciálegyenletnek nevezzük, mert az ismeretlen függvényen kívül a deriváltja is szerepel az egyenletben. A megoldás azonban - mint látni fogjuk - nem lesz egyértelmű, hiszen a növekedés változásából a népesség számra nem tudok következtetni. Meg kell adni a kiinduló állapotot, pl. az $N(0) = n$ értéket.

Mielőtt megoldanánk az egyenletet, finomítsuk a modellünket. Az előbbi egyenlet azt fejezi ki, hogy a társadalom bizonyos százaléka elpusztul, és bizonyos százalékban reprodukálódik, a két százalékszám különbsége a fenti c szám. Tegyük fel most, hogy az eltelt idővel arányosan bizonyos mennyiségű bevándorlással is kell számolni. Akkor a kiinduló egyenletünk

$$N(t+h) - N(t) = cN(t)h + ah$$

alakra módosul, amiből ugyanezzel az eljárással az

$$N'(t) = cN(t) + a$$

differenciálegyenlet adódik (c is és a is lehet negatív szám is).

Az első egyenlet megoldása egyszerű. $N(t)$ -vel elosztva mindkét oldalt

$$\frac{N'(t)}{N(t)} = c$$

adódik, ahol a baloldalon $\frac{N'(t)}{N(t)} = (\ln N(t))'$, vagyis az $\ln N(t)$ függvény deriváltja c , így

$\ln N(t) = ct + C$, amiből $N(t) = e^{ct+C}$. Itt kell felhasználni, hogy $N(0) = n$, amiből $e^C = n$ következik, a megoldás tehát $N(t) = ne^{ct}$.

A második egyenlet megoldását később tárgyaljuk.

20.2. A differenciálegyenletek osztályozása

Több szempont alapján lehet osztályokba sorolni a differenciálegyenleteket. Legelső szempont a differenciálás. Ha többváltozós függvény az ismeretlen függvény és a parciális deriváltakat tartalmazza az egyenlet, akkor parciális differenciálegyenletekről beszélünk. Egy változós esetben - ha a parciálissal szembeállítjuk - akkor közönséges differenciálegyenletekről beszélünk.

A differenciálegyenlet elsőrendű, ha az ismeretlen függvénynek csak az első deriváltja (és esetleg maga a függvény) szerepel az egyenletben. Magasabbrendű, ha magasabb deriváltja is szerepel benne; így másodrendű egyenlet tartalmazza magán a függvényen kívül az első és a második deriváltját. Hasonló értelemben beszélhetünk harmadrendű, negyedrendű, ... differenciálegyenletekről. Itt csak elsőrendű közönséges differenciálegyenletekről lesz szó.

Az elsőrendű differenciálegyenlet lineáris, ha $y(x)$ -szel jelölve az ismeretlen függvényt - a 0-ra redukált egyenlet baloldala y, y' lineáris függvénye, a lineáris függvényben fellépő együtthatók viszont x -nek tetszőleges adott (nem feltétlenül lineáris) függvényei. Az elsőrendű lineáris differenciálegyenlet általános alakja:

$$a(x)y' + b(x)y = c(x).$$

Ha az együtthatók állandók, azaz $a(x) = 1, b(x) = b, c(x) = c$, akkor állandó együtthatós elsőrendű lineáris differenciálegyenletről beszélünk.

20.3. Elsőrendű lineáris differenciálegyenlet megoldása

Először, didaktikai okokból, az állandó együtthatós elsőrendű lineáris differenciálegyenletet oldjuk meg. Az általános alakja:

$$y' + by = c.$$

Szorozzuk be mindkét oldalt e^{bx} -nel, akkor a baloldal ye^{bx} deriváltja:

$$y'e^{bx} + by e^{bx} = (y e^{bx})' = c e^{bx}.$$

$y e^{bx}$ tehát $c e^{bx}$ primitív függvénye, azaz $\frac{c}{b} e^{bx} + C$, innen

$$y = \frac{c}{b} + C e^{-bx},$$

ahol a C együtthatót a kezdeti feltétel alapján kell meghatározni.

Általános esetben hasonló az eljárás. Az egyenlet együtthatói most függvények, az általános alak:

$$y' + b(x)y = c(x).$$

Számoljuk ki $b(x)$ primitív függvényét (elég egy verziót megadni, így a C -t itt nem kell figyelembe venni), legyen ez $B(x)$. Szorozzuk be mindkét oldalt $e^{B(x)}$ -nel, akkor a bal oldal $y e^{B(x)}$ deriváltja lesz:

$$y' e^{B(x)} + y b(x) e^{B(x)} = (y e^{B(x)})' = c(x) e^{B(x)},$$

amiből - az előzőekhez hasonlóan -

$$y e^{B(x)} = \int c(x) e^{B(x)} dx + C,$$

$$y = e^{-B(x)} \int c(x) e^{B(x)} dx + C e^{-B(x)},$$

ami az általános megoldást jelenti. Konkrét esetben a C -t a kezdeti feltételekből kell meghatározni.

20.4. Szétválasztható differenciálegyenlet

A szétválasztható differenciálegyenlet általános alakja:

$$y' = f(y)g(x).$$

Rendezzük át

$$\frac{y'}{f(y)} = g(x)$$

alakba, majd integráljuk mindkét oldalt x szerint:

$$\int \frac{y'(x)}{f(y(x))} dx = \int g(x) dx + C.$$

A baloldali integrál $z = y(x)$ helyettesítéssel, amikor is $dz = y' dx$, az

$$\int \frac{y'(x)}{f(y(x))} dx = \int \frac{dz}{f(z)}$$

alakot ölti. Ha $F(z)$ -vel jelöljük $\frac{1}{f(z)}$ primitív függvényét, akkor

$$\int \frac{y'(x)}{f(y(x))} dx = \int \frac{dz}{f(z)} = F(z) = F(y(x)),$$

azaz

$$F(y(x)) = \int g(x) dx,$$

ami az integrálás elvégzése után implicit alakban megadja az $y(x)$ függvényt.